

Airbnb New User Bookings : Problem Description

Problem statement :

Airbnb is an online marketplace and hospitality service, enabling people to lease or rent short-term lodging including vacation rentals , apartment rentals, homestays , hostels beds, or hotel rooms.

New users on Airbnb can book a place to stay in 34,000+ cities across 190+ countries. By accurately predicting where a new user will book their first travel experience, Airbnb can share more personalized content with their community, decrease the average time to first booking, and better forecast demand. We need to predict the first travel destination of a new user based on his personalized content .

Data Description :

Data type : csv files

- age_gender_bkts.csv - summary statistics of users' age group, gender, country of destination
- countries.csv - summary statistics of destination countries in this dataset and their locations
- sample_submission.csv - correct format for submitting your predictions
- Sessions.csv - (user_id , action , action_type , action_detail , device_type , secs_elapsed)
- test_users.csv - (id , date_account_created , timestamp_first_active , date_account_created, date_first_booking , gender , age , signup_method , signup_flow , language , affiliate_channel , first_affiliate_tracked , signup_app , first_device_type , first_browser , country_destination)
- train_users.csv - (Similar to test data)

Data size : 664.2 MB

Project reports :

<https://cseweb.ucsd.edu/~jmcauley/cse255/reports/fa15/045.pdf>

<https://cseweb.ucsd.edu/~jmcauley/cse190/reports/fa15/027.pdf>

<https://cseweb.ucsd.edu/~jmcauley/cse190/reports/fa15/006.pdf>

<https://cseweb.ucsd.edu/classes/wi17/cse258-a/reports/a065.pdf>

Useful Links :

<http://blog.kaggle.com/2016/03/17/airbnb-new-user-bookings-winners-interview-2nd-place-keiichi-kuroyanagi-keiku/>

<http://blog.kaggle.com/2016/03/07/airbnb-new-user-bookings-winners-interview-3rd-place-sandro-vega-pons/>

<http://rohanrao91.blogspot.in/2016/02/airbnb-new-user-bookings.html>

Algorithms or Techniques :

1. Exploratory data analysis
2. Naive Bayes algorithm
3. Random forests
4. Gradient boosting Decision trees
5. K Nearest neighbour algorithm
6. Stacking
7. Feature Selection